Home advantage in speed skating

Koning, Ruud H.

Home Advantage in Speed Skating: Evidence from Individual Data

Ruud H. Koning[1]

SOM Research Report 03F38

[1]Department of Econometrics, Post Box 800, 9700 AV Groningen, The Netherlands, email: r.h.koning@eco.rug.nl.

Abstract

Home advantage is a well documented phenomena in many types of sports. Home advantage has been shown to exist for team sports (soccer, hockey, football, baseball, basketball) and for countries organizing sports tournaments like the Olympics and World Cup Soccer. There is also some evidence for home advantage in some individual sports, but there is much less literature available in that case.

In this paper we address the issue of home advantage in speed skating. From a methodological point of view, it is difficult to identify home advantage, because skaters vary in their abilities, and the conditions of different tournaments vary.

We establish the existence of a small, but significant home advantage, using a generalized linear mixed model, with random effects for skaters, and fixed effects for skating rinks and seasons. Even though the home advantage effect exists, it is very small when compared to variation in skating times due to differences of rinks and individual abilities.

Keywords: home advantage, speed skating, generalized linear mixed model, random effects.

1 INTRODUCTION

Home advantage is a well documented feature in many types of sports. By now a significant volume of research has been published documenting its existence, and quantifying its effect on the outcome of sports contests. Home advantage has been shown to exist for individual sports (alpine skiing), team sports (soccer, hockey, football, baseball, basketball), and for countries organizing sports tournaments like the Olympics and World Cup Soccer. Literature on home advantage in sports is reviewed in Nevill and Holder (1999).

Home advantage is usually attributed to four different factors: crowd support, familiarity with local conditions, reduced travel time for home athletes, and, finally, the rule factor (Nevill and Holder (1999)). These factors are of different relevance for different sports, and we examine them later in detail for our case at hand.

In this paper, we test for the existence of home advantage in speed skating. Also, we compare the magnitude of home advantage, if any, to other sources of variation of skating times. To our knowledge, this has not done before. The case of speed skating is interesting, for a couple of reasons. First of all, it is an individual sport. Most research of home advantage has concerned team sports, Bray and Carron (1993) and Holder and Nevill (1997) being notable exceptions. Bray and Carron prove the existence of home advantage in alpine skiing, and Holder and Nevill find only little evidence in tennis and golf. Individual results are, more often than not, not measured in a full competition and that complicates estimation of the home advantage effect. Second, skating performances are rated with an absolute measure of performance: the time skated, so performances are rated on a ratio scale and not on some ordinal scale. Third, we have a unique data set with multiple measurements per skater within and between seasons, and skaters skate different distances. This allows us to identify home advantage without imposing very strong parametric assumptions on the statistical model, while at the same time allowing for different abilities of skaters.

The plan of this paper is as follows. In section 2 we discuss speed skating and the possible role of home advantage. A statistical model of speed skating results, and identification and estimation of a home advantage effect is the topic of section 3. We end with conclusions in section 4.

2 Speed Skating and Home Advantage

Skating is a winter sport with a long history. In fact, the International Skating Union was founded in 1893 and is the oldest governing international winter sport federation. Continuous progress has improved the results of skaters. The first world record on 1500m was skated on a frozen lake near Groningen (The Netherlands) by a skater wearing woolen clothes. The last world record on this distance was skated in the high-tech skating rink of Salt Lake City, which is on a high altitude and has a covered roof[1]. Moreover, skating suits nowadays are aerodynamic, and the skates have improved impressively as well. A concise history of skating, and especially a discussion of the role of technological progress, can be found in Kuper and Sterken (2003b).

Modern speed skating usually takes place on refrigerated ovals. Currently, only a few non-refrigerated ovals are used for international competitions. Skaters start at the the shot of a starting gun (which starts the time measurement) and skate 400m laps in one direction. Every lap they change from the inside lane to the outside lane, and vice versa. The time skated is measured when they pass the finish line crossing a laser beam (this equipment is also used to measure the time for intermediate laps).

Skaters compete against each other during tournaments or events. Some of these events are national events (the Dutch or Norwegian championship), and because within a country only a few skating rinks exist, it is not reasonable to search for home advantage for a particular skater within a country. Variation of performances will be due to variation of the quality of the skaters and the different rinks in a country, not to any perceived home advantage. However, skaters also meet at international competitions. The most important international competitions are the World Cup (a series of meetings in different venues, culminating in the World Cup Final), the World Championship Distances, the World Championship Sprint, the European and World Championship all round, and the Olympic Winter Games. These international meetings are held in different places all over the world and therefor provide a suitable basis for examining the existence of home advantage. In this paper, we do not use observations on any all round tournament, where skaters have to skate four distances in two or three days. These tournaments are rather different from contests where a skater can focus on one or two distances. Therefor, we use only observations on World Cup meetings, World Championship Distances, and Olympic Winter Games. Distances skated by men are 500m, 1000m, 1500m, 5000m, and 10000m. Women skate 500m, 1000m, 1500m,

[1]The special quality of the Salt Lake City skating rink is examined in detail by Reese (2003).

3000m, and 5000m. Both in the case of men and women, the longest distance is highly specialized with only a few skaters competing for top positions.

Since World Cup meetings, the World Championship Distances, and the Olympic Winter Games are held in different countries, there may be some identifiable home advantage for skaters of the country hosting the event. Nevill and Holder (1999) discuss the factors contributing to home advantage in detail. They distinguish between four factors: crowd support, familiarity with local conditions, reduced travel time for home athletes, and, finally, the rule factor. To what extend are these factors applicable to speed skating?

Crowd support can help an athlete to perform better than anticipated. In skating, crowds tend to cheer for most if not all skaters. Antagonism between crowds and skaters is rare. If crowd support is a factor in the performance of skaters at all, it does not discriminate much between home skaters and skaters from other countries. Crowd support could, in theory, influence decisions by referees. Referees in skating have to make two types of decisions: whether or not the skater moved before the shot of the starting gun, and whether or not the skater from the inner lane gave way to the skater coming from the outer lane when they change lanes. Both types of decision are usually uncontroversial and no skaters are disqualified during an event. The most frequent cause of skater not finishing is not the decision of a referee, but a fall during the race.

Familiarity with local conditions is potentially a more important factor in determining home advantage. Skating rinks lie at different altitudes, and some are covered while others are not. Also, there is a difference in curvature of different rinks. Some ovals have tighter bends than other ovals. A skater who is used to training frequently at a high speed skating rink (for example, the Olympic Oval in Calgary, Canada) may develop skills that are lacked by skaters training at lower speed rinks. Especially skating bends at high speed is difficult and requires a lot of practice. Such techniques are best learnt at these high speed skating rinks, which gives an advantage to skaters who can train on such rinks regularly. For this reason, skaters from countries at low altitude have training sessions at such high speed rinks in the summer or early in the season. According to skating folklore, 'power skaters' perform well at covered and uncovered low altitude skating rinks, while 'flyers' are the best performers under the controlled conditions of high altitude covered rinks. Some anecdotal evidence for this assertion can be found in the fact that the correlations of skating times between different rinks[2]

[2]To be precise, 1500m times in the 2001/2002 season.

tend to be higher between similar rinks than between dissimilar rinks.

Reduced travel time for home athletes is unlikely to be an important factor causing home advantage. During the season, most skaters participate in the same events and hence, they are subject to the same time differences. Skating events take place in the weekends, so the rest of the week can be used for travel and overcoming travel fatigue.

Finally, it is unlikely that the rules benefit home skaters. Rules can be manipulated to benefit a skater in two ways. First, the ice is cleaned at regular intervals, and there is a small advantage to skating on a clean rink. However, the decision when to clean the ice is made by the referee(s) and a representative of the International Skating Union. This decision is communicated to the coaches *before* the order of pairs of skaters is known. During the event, it may be decided that more frequent treatments of the rink are necessary and at this decision to change the order of cleaning the rink could benefit some skaters and harm other skaters. The second important issue is the order of skating. Most skaters prefer to skate last, so that they know the times of their competitors. The order of skating is determined as follows. Skaters are allocated to groups based on their times in previous events, or on their performance on other distances during the same event. The order and pairing is then determined by draw. This leaves little room for manipulation.

Summarizing, we see two possible sources of home advantage in skating: crowd support and familiarity with the track and skating at that particular altitude. Whether these sources of home advantage are statistically significant or empirically relevant in determining the outcome of a skating contest is not clear, though. Balmer, Nevill, and Williams (2003) compare different sports examining home advantage in summer Olympic Games. They find that the role of referees and jury's are important sources of home advantage. Considering the very marginal importance of referee decisions in skating contests, we expect home advantage, if any, to be small.

3 An Empirical Model of Speed Skating Results

In this section we proceed to measure home advantage in speed skating, and test whether it statistically significant. We use a database with finishing times of participants in World Cup meetings, World Championship Distances, and Olympic Winter Games from 1986/87 to 2002/03. The data set has observations both on men and women skaters. A detailed description of the data set and selections to clean the data set are given in Appendix A. Here, let it suffice to note that we have data on 17 seasons, 21963 times of men, and 15905 times of women skaters. All skating times were transformed to their 500m equivalent, a common way of making different distances comparable. This speed per 500m is our measure of performance and is going to be used to assess the existence and magnitude of home advantage. From now on, we omit the dimension of our measure of performance, so when we refer to ‘speed’ or ‘time’ we mean ‘time (in seconds) per 500m’.

Most empirical studies of home advantage consider either team data, or national aggregate data. Usually, proof of home advantage is found in a win percentage that exceeds 50%. This assumes that the home team (or country) and the opposing teams are of equal quality, an assumption that is not realistic in many cases. A better approach is to separate home advantage and team quality, for example along the lines of Clarke and Norman (1995). Their model is applied to soccer results, and home advantage is measured as the expected difference of goals scored and conceded when playing a (hypothetical) opponent of the same quality as the home team. Another advantage of their approach is that the parameters of the model can be estimated both based on a full competition and a partially completed competition schedule. Our data are unbalanced in the sense that there will not be a home advantage for every skater, because skating tournaments take place in a few countries only. Moreover, we encounter an additional complication. Performances of skaters are not comparable between different rinks and in this respect our study differs significantly from all other studies of home advantage. In figure 1 we graph the distribution of skating speed by means of a box-and-whisker plot[3], conditioning on sex of the skater, and cover and altitude of the rink. Also, we distinguish between home results and away results. Each cell of figure 1 is based on 17 seasons of data. Clearly, the median speed varies by these conditioning factors and the variation of speeds caused by these observable factors is bigger than any variation due to home advantage. However, it is apparent from figure 1 that there is some home advantage, mainly for covered skating rinks. A possible explanation for this may be that weather conditions vary during outdoor events, masking any home advantage effect. Covered skating rinks give better measurements of the quality of a skater.

Another issue we need to take into account is improvement of skating speeds over time. In figure 2 we graph the median speeds by season for both men and women,

[3]In these box-and-whisker plots, the left side of the box indicates the first quartile, and the right side the third quartile. The dot in each box is the median. The whiskers indicate the minimum and maximum, except when these deviate more than 1.5 times the interquartile range. Observations deviating more than 1.5 times the interquartile range are depicted by dots.

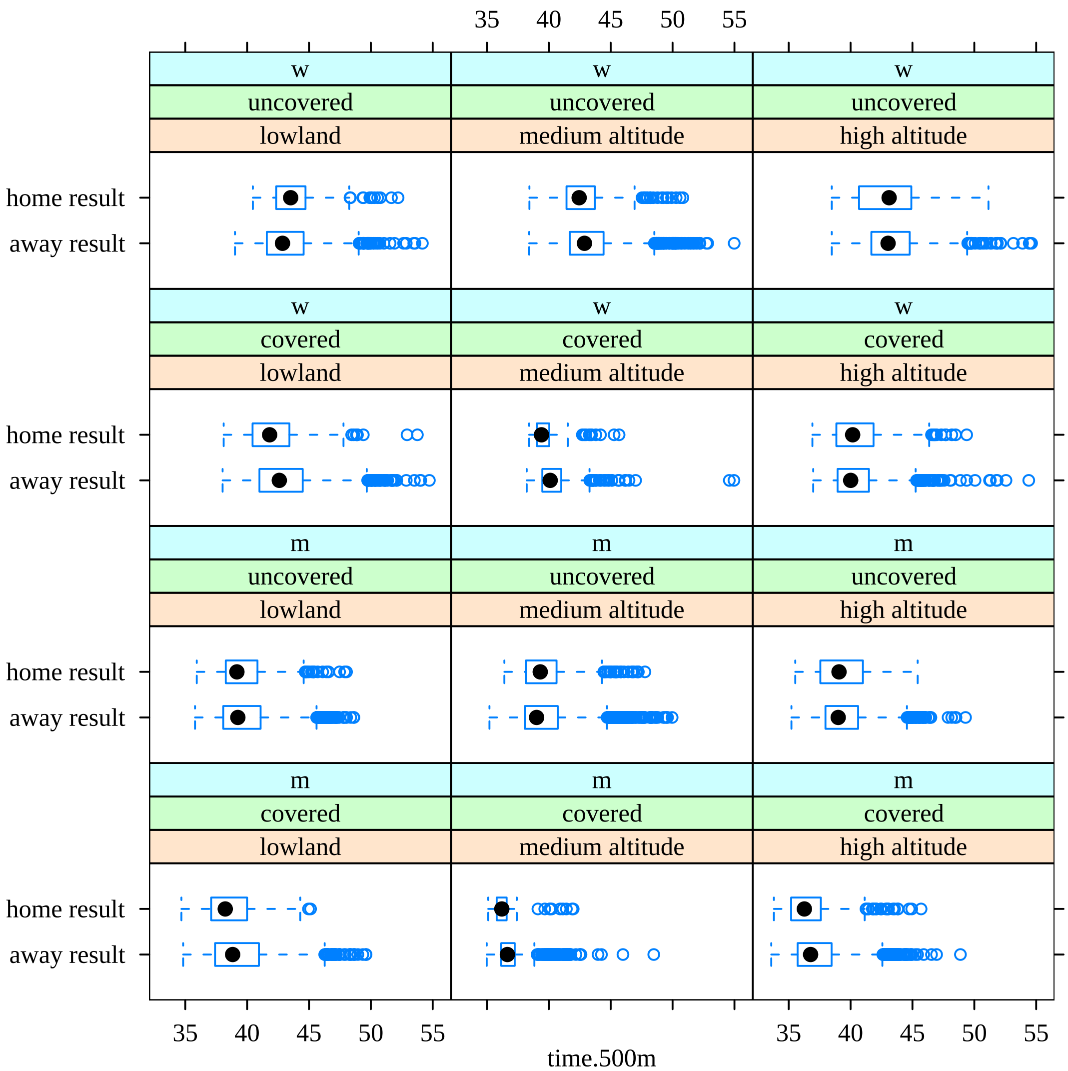


Figure 1: Distribution of speed by sex, cover, and altitude.

conditional on the distance. For every distance we see a marked improvement over time. Of course, part of this improvement is due to the advent of covered skating rinks, but also other factors have contributed. The most significant example of technological progress is the introduction of the klap skate in the 96/97 season (see also Kuper and Sterken (2003b)). Other important examples are improvement in training methods (Gemser and De Koning (2001)) and improvements of clothing (Kuper and Sterken (2003a)).

We have by now identified three important sources of variation of skating performances: covering of the track, altitude of the track, and general improvement of skating speeds over time. Figure 1 suggests that besides these factors home advantage may also be a determinant. After this description of the general features of our data set, we turn to the identification and estimation of home advantage.

So far, we have not been explicit about our exact definition of home advantage in the context of speed skating. Home advantage is usually defined as the performance advantage of an athlete, team, or country when they compete at a home ground. This definition is not very fruitful in our case at hand: from figure 1 is is clear that variation of skating times due to home advantage is small when compared to the variation caused by differences of altitude and cover of tracks. Skating conditions are not comparable between countries, fast speed high altitude skating rinks are found only in Canada (Calgary) and the US (Salt Lake City) while only lowland rinks are available in The Netherlands. Because of this, we opt for a different definition of home advantage: home advantage is the performance advantage of an athlete, team, or country when they compete at a home ground compared to their performance under similar conditions at an away ground. The controlled experiment implicit in this definition is not observed. Observed skating times vary by season, distance, and rink. We allow for this variation by using a statistical model for the expected skating time, given relevant covariates one of which is home advantage.

An important issue in identification of home advantage is that one has to allow for quality variation of the skaters, as has been pointed out in, for example, Balmer, Nevill, and Williams (2001). Suppose for example that US skaters are much better than all other skaters. Any good performance of these US skaters during a meeting in the US should be attributed then to their superior quality, and not to home advantage. We are able to allow for quality variation because our data set contains multiple results by the same skater, and that enables us to estimate his (or her) quality.

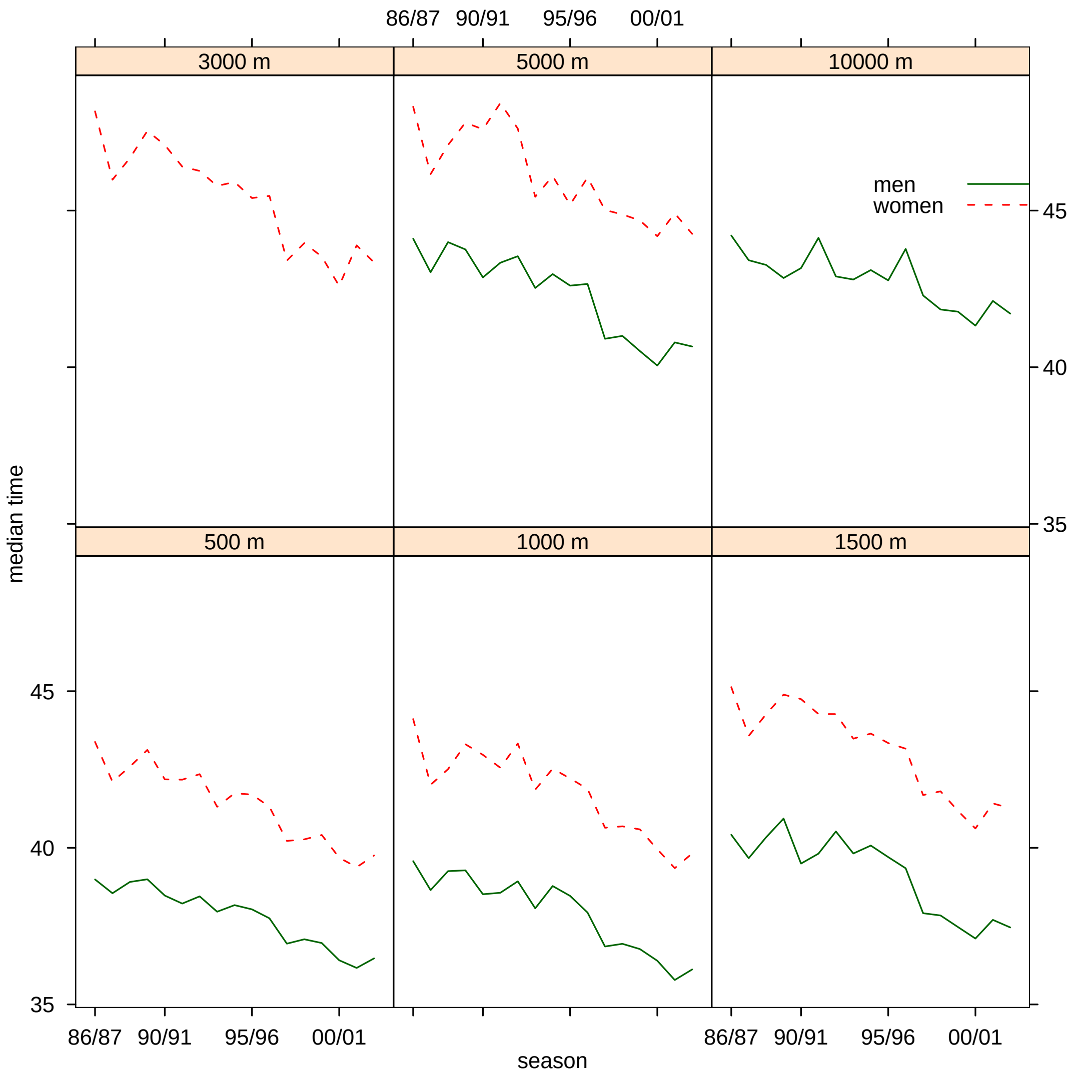


Figure 2: Development of median times by season for men (solid line) and women (dashed line).

Let the speed of skater i at distance k during event t be denoted by T_{ikt}. Speed depends on the rink, the distance skated, and the season when event t takes place. Even though the trends displayed in figure 2 suggest a linear improvement of performance over time, we are hesitant to impose such a linear effect. A linear time trend does not allow sudden improvements of skating times due to important innovations like the klapskate. Besides these determinants of skating speed, we also allow for the possibility that speeds skated during special events as the World Championship Distances or the Olympic Games are higher than expected (when compared to speeds skated at World Cup meetings). Finally, we allow for home advantage.

We model skating times using a generalized linear mixed model, with a logarithmic link function. A somewhat similar approach to modelling skating times is taken by Reese (2003) who uses a hierarchical model (with more levels than one) to model skating times. The distribution is assumed to be a Gamma distribution, because skating times are skewed to the right. The specification for the logarithm of the conditional mean is:

$$\log \mathcal{E} T_{ikt}|\alpha_i = \sum_{r=1}^{44} \beta_r R_{rt} + \sum_{k=1}^{5} \gamma_k D_{ikt} + \delta H_{it} + \theta_1 OG_t + \theta_2 WSDCh_t + \sum_{\tau=1}^{16} \phi_\tau S_{\tau t} + \alpha_i \tag{1}$$

The β-parameters are the fixed effects, one for each skating rink. The γ-parameters quantify the effect of distance: average speed decreases when the observation is from a longer distance. The variables D_{ikt} are dummy variables, with $D_{i1t} = 1$ for a 1000m event. D_{i2t}, D_{i3t}, D_{i4t}, and D_{i5t} correspond to 1500m, 3000m, 5000m, and 10000m events respectively. Note that γ_3 cannot be estimated for men because they don't skate 3000m frequently in official events. Since women do not skate 10000m, γ_5 is not estimable for them. As distance is a categorical variable, we had to choose a reference category which is 500m. The δ parameter measures home advantage. If home advantage exists, we expect δ to be negative. The θ parameters measure tournament effects: θ_1 captures the effect of Olympic Games, and θ_2 the effect of World Championship Distances. A World Cup meeting is the reference category. ϕ_τ are the fixed seasonal effects. A season dummy is included for each of the seasons 1987/88 to 2002/03. The 1986/87 season is the reference category.

The specification as discussed so far does not allow for quality variation between skaters. This effect is captured by α_i. We have observations on 665 male skaters, and 453 female skaters. Clearly, it is not feasible to estimate a separate quality parameter for each of them. Instead, we model quality variation by assuming that deviations from average quality follow a normal distribution with mean 0 and variance σ_α^2. The parameter α_i is fixed for each skater between different events and seasons.

The model was estimated using penalized quasi maximum likelihood, see Breslow and Clayton (1993) and Venables and Ripley (2002). The venue specific fixed effects $\beta_r, r = 1, \ldots, 44$ are given in table 2 in appendix B. The fixed effects are summarized graphically in figure 3, conditioning again on cover and altitude. Clearly, men skate faster than women on each type of track (the average difference is 3.7 seconds), and skaters are faster on covered rinks than uncovered rinks (the average difference is 0.874 seconds for men and 0.934 seconds for women). Note also that there is some variation of the fixed effects in most cells, which justifies our approach to estimate separate fixed effects for each rink.

In table 1 we give the estimation results of model (1). Point estimates and their standard errors are provided. For ease of interpretation, we also give the exponentiated point estimate, which is the factor of proportionality of the effect[4]. For example, 1.003 (the exponentiated coefficient of 1000m for men) should be interpreted as follows. Keeping all other factors (including the random individual effect) constant, the 1000m speed for men is 0.3% slower than the 500m speed for men. Using similar reasoning, the expected speed on 1500m is 2.6% slower than the expected speed on the 500m distance.

The estimates for both men and women are consistent to prior expectations. Speed is an decreasing function of distance, as the coefficients of the distances show. Time skated at a 1000m event is only marginally less than time skated on a 500m event (0.3% for men and 0.9% for women). The reason for this that the initial speed at any event is zero, and it takes a little bit of time to start. On longer distances, this loss of time is averaged over a longer distance and hence improves average speed. On the other hand, fatigue sets in on longer distances, which decreases speed. Apparently, the latter effect dominates for distances of 1500m and more.

We also see that both men and women skate significantly faster at the major tournaments: the Olympic Games and the World Distances Championship. The effect is of

[4]This is a correct interpretation of the coefficients, because the logarithm of the expected skating times is modelled, and not the expectation of the logarithm of skating times.

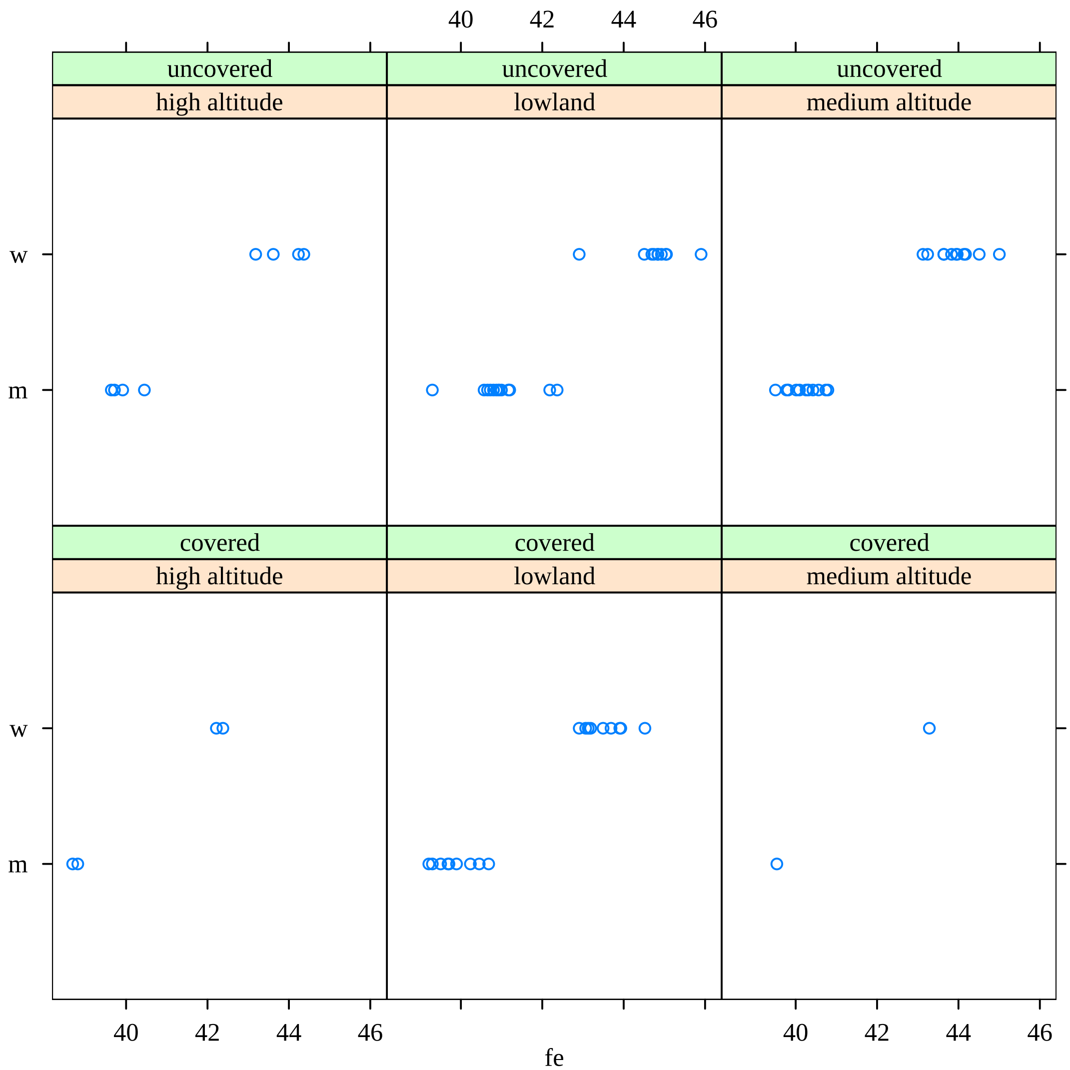


Figure 3: Distribution of rink fixed effects, by cover and altitude.

Table 1: Estimation results for men and women.

	Men			Women		
	estimate	st.err.	exp(est.)	estimate	st.err.	exp(est.)
distance1000 m	0.0026	0.0004	1.003	0.0089	0.0004	1.009
distance1500 m	0.0259	0.0005	1.026	0.0348	0.0006	1.035
distance3000 m				0.0829	0.0007	1.086
distance5000 m	0.1019	0.0006	1.107	0.1140	0.0010	1.121
distance10000 m	0.1326	0.0010	1.142			
home result	−0.0020	0.0005	0.998	−0.0025	0.0006	0.997
OG	−0.0100	0.0009	0.990	−0.0119	0.0010	0.988
WSDCh	−0.0099	0.0008	0.990	−0.0100	0.0008	0.990
season87/88	−0.0048	0.0011	0.995	−0.0051	0.0017	0.995
season88/89	−0.0098	0.0011	0.990	−0.0008	0.0019	0.999
season89/90	−0.0067	0.0012	0.993	0.0034	0.0018	1.003
season90/91	−0.0120	0.0013	0.988	0.0012	0.0019	1.001
season91/92	−0.0121	0.0013	0.988	−0.0048	0.0020	0.995
season92/93	−0.0135	0.0013	0.987	−0.0091	0.0019	0.991
season93/94	−0.0130	0.0013	0.987	−0.0130	0.0019	0.987
season94/95	−0.0154	0.0013	0.985	−0.0154	0.0019	0.985
season95/96	−0.0221	0.0013	0.978	−0.0196	0.0019	0.981
season96/97	−0.0256	0.0013	0.975	−0.0274	0.0019	0.973
season97/98	−0.0459	0.0013	0.955	−0.0500	0.0019	0.951
season98/99	−0.0567	0.0013	0.945	−0.0572	0.0020	0.944
season99/00	−0.0619	0.0013	0.940	−0.0618	0.0019	0.940
season00/01	−0.0681	0.0013	0.934	−0.0711	0.0019	0.931
season01/02	−0.0677	0.0013	0.935	−0.0704	0.0019	0.932
season02/03	−0.0684	0.0014	0.934	−0.0729	0.0020	0.930
σ_α	0.0308			0.0335		

the order of 1%, both for men and for women for both types of events. The existence of this effect could be caused by national selection of skaters participating in these tournaments: only the very best of the best qualify for such events. Also, these tournaments attract a lot of media coverage, which makes it important for skaters to perform as good as they can. Miserable performance during these two tournaments can have dire consequences for future sponsor deals and access to training facilities and new materials. Qualification criteria for World Cup meetings are usually less strict.

General progress is significant, and has affected women skaters slightly more than men skaters. During the 17 years in our data set, average speed has increased by 7% for women and 6.6% for men. Note the improvement of approximately 2% in the 1997/98 season. This corresponds to the broad acceptation of the klapskate (see Kuper and Sterken (2003b)).

Now consider our estimate of home advantage. Home advantage is both positive, and statistically significant. The point estimates are approximately 0.2%, both for men and for women. Even though home advantage is significant, the effect is very small compared to variation of times caused by other factors like general progress and variation between skating rinks. It may be argued that home advantage is confounded with access to new technology, if skaters from some countries have earlier access to new technology than other skaters. Two examples come to mind: the Dutch female skaters started using the klap-skate in the 1996/97 season, earlier than other skaters. American and Dutch skaters used special suits in the 2001/02 season, which had a positive impact on their skating times (Kuper and Sterken (2003a)). Because not all skaters were able to use these new advances, this progress is not fully captured by the season fixed effects. For that reason we re-estimated model (1), removing those seasons from the analysis. The new point estimates for home advantage are within one standard error of those reported in table 1. It seems that our estimate of home advantage is not confounded with early access to new technologies by a select group of skaters.

Finally, we come to the standard deviation of the individual effects α_i. The standard deviation of the individual effects is 3.1% (for men) and 3.4% (for women). This suggests that variation of abilities of different skaters is empirically important. A 95% confidence interval for individual effects of male skaters is $(-0.0604, 0.0604)$, and a corresponding interval for women is $(-0.0657, 0.0657)$. A comparison of these intervals to the estimated home advantage effect shows that the latter is dwarfed by variation in individual abilities. This is also confirmed by the signs and magnitudes of the estimates of the best linear predictors of the individual effects, as some examples show: Koss 0.962, Romme 0.958, Niemann-Stirnemann 0.939, and LeMay-Doan 0.956. These skaters are well known to be exceptional and an estimated individual effect smaller than 1 indicates that they skate faster than an average skater under similar circumstances.

It is interesting to note that the standard deviation of the random effects is larger for women than for men. This suggests that there is more competition in men events than in women events. This observation is in accordance with Gould's hypothesis that quality variation of athletes decreases over time when a sport matures (Gould (1997)). International speed skating (as analyzed in this paper) for men has a longer history than speed skating by women, and therefor we would expect less variation of abilities of top athletes.

4 Summary and Conclusions

In this paper we have examined whether there is any proof of the existence of home advantage in speed skating. Performance in speed skating can be measured objectively by the time taken to cover a certain distance, and referee decisions are usually of no influence on the results. According to the existing literature, these two considerations indicate that it is unlikely that home advantage exists.

There are problem in estimating home advantage: skaters differ in their abilities, and some skating rinks are of better quality than others. To accommodate these problems, we used a statistical model that allows for quality variation of skating rinks, and for different abilities of skaters. Using this model, we identified a significant home advantage, that is similar for men and women: it is approximately 0.2%. However, the magnitude of this home advantage effect is very small, when we compare it to other sources of variation like variation of individual abilities, variation of quality of skating rinks, and improvement of skating times over time. Hence, we conclude that even though a significant home advantage exists, it plays a very minor role in determining the performance of skaters. In this sense, skating differs from many other sports where a well established non-negligible home advantage exists.

ACKNOWLEDGEMENTS

The data used in this paper were kindly provided by Jeroen Heijmans (`http://weasel.student.utwente.nl/~speedskating/`). The statistical analysis was performed using the *R* statistical software system, Ihaka and Gentleman (1996), which can be obtained from `www.r-project.org`. The original paper was typeset using mikTEX, which can be obtained from `www.miktex.org`. I thank Gerard Kuper for detailed comments on an earlier draft of this paper.

REFERENCES

Balmer, N.J., A.M. Nevill, and A.M. Williams (2001). Home advantage in the Winter Olympics (1908–1998). *Journal of Sports Sciences 19*(2), 129–139.

Balmer, N.J., A.M. Nevill, and A.M. Williams (2003). Modelling home advantage in the Summer Olympic Games. *Journal of Sports Sciences 21*(6), 469–478.

Bray, S.R. and A.V. Carron (1993). The home advantage in alpine skiing. *Australian Journal of Science and Medicine in Sport 25*, 76–81.

Breslow, N.E. and D.G. Clayton (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association 88*(421), 9–25.

Clarke, S.R. and J.M. Norman (1995). Home ground advantage of individual clubs in English soccer. *The Statistician 44*(4), 509–521.

Gemser, H. and J. De Koning (2001). Schaatsen met de computer: Op zoek naar prestatiebepalende factoren (skating with the computer: Searching for perfomanace enhancing factors). Presentation at the symposium ‘Sport and Computer’, Amsterdam.

Gould, S.J. (1997). *Full House*. New York: Random House.

Holder, R.L. and A.M. Nevill (1997). Modelling performance at international tennis and golf tournaments: Is there a home advantage? *The Statistician 46*(4), 551–559.

Ihaka, R. and R. Gentleman (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics 5*(3), 299–314.

Kuper, G.H. and E. Sterken (2003a). Do skin suits increase average skating speed? Manuscript, Department of Economics, University of Groningen.

Kuper, G.H. and E. Sterken (2003b). Endurance in speed skating: The development of world records. *European Journal of Operational Research 148*(2), 293–301.

Nevill, A.M. and R.L. Holder (1999). Home advantage in sport: An overview of studies on the advantage of playing at home. *Sports Medicine 28*(4), 221–236.

Reese, C.S. (2003). Is Salt Lake City's ice the fastest on earth?: A statistical investigation. Paper presented at the Joint Statistical Meeting, San Fransisco.

Venables, W.N. and B.D. Ripley (2002). *Modern Applied Statistics with S* (Fourth ed.). New York: Springer.

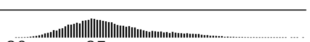
A Additional Information on the Data Set

In this appendix we provide some more information on the data set. The data set contains all results on World Cup meetings, World Championship Distances, and Olympic Winter Games, starting 1986/87 season and ending with the 2002/03 season. All times were recalculated as time per 500m, a common way of comparing times between different distances. Some skaters fell during their race, but they finished their race anyway. This results in a uncharacteristically high time per 500m, so we removed observations which took longer than 50 seconds per 500m (men), and 55 seconds per 500m (women). Also, we removed all 3000m observations for men because this distance is non-standard. After these selections, we have 21963 observations for men, and 15905 observations for women. The distribution over distances and seasons is given in figures 4 and 5. The width of each column is proportional to the number of observations for that season. Both graphs show that the number of observations per distance does not vary that much. Moreover, the number of observations of female skaters has increased since the 1994/95 season. There are only few observations on the longest distances 10000m for men and 5000m for women.

Men

7 Variables 21963 Observations

time.500m



n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
21963	0	6168	39.15	35.78	36.34	37.39	38.69	40.53	42.77	43.89

```
lowest : 33.59 33.66 33.77 33.81 33.86
highest: 49.53 49.59 49.60 49.61 49.96
```

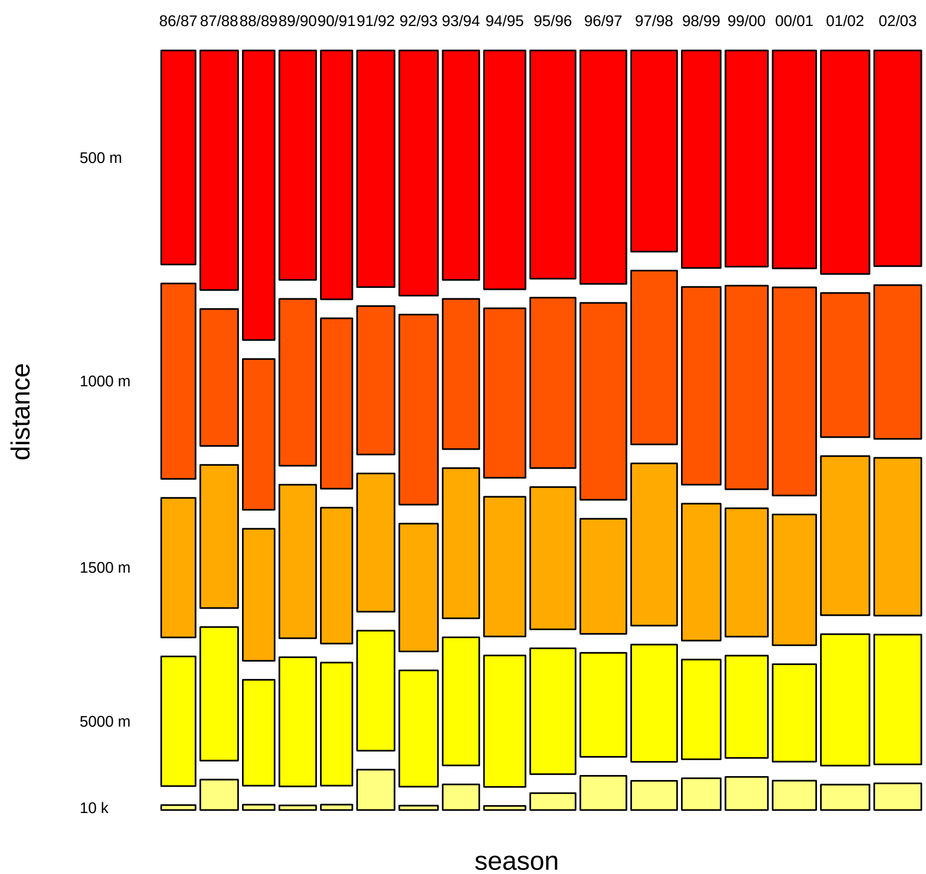


Figure 4: Distribution of observations by distance and season (men).

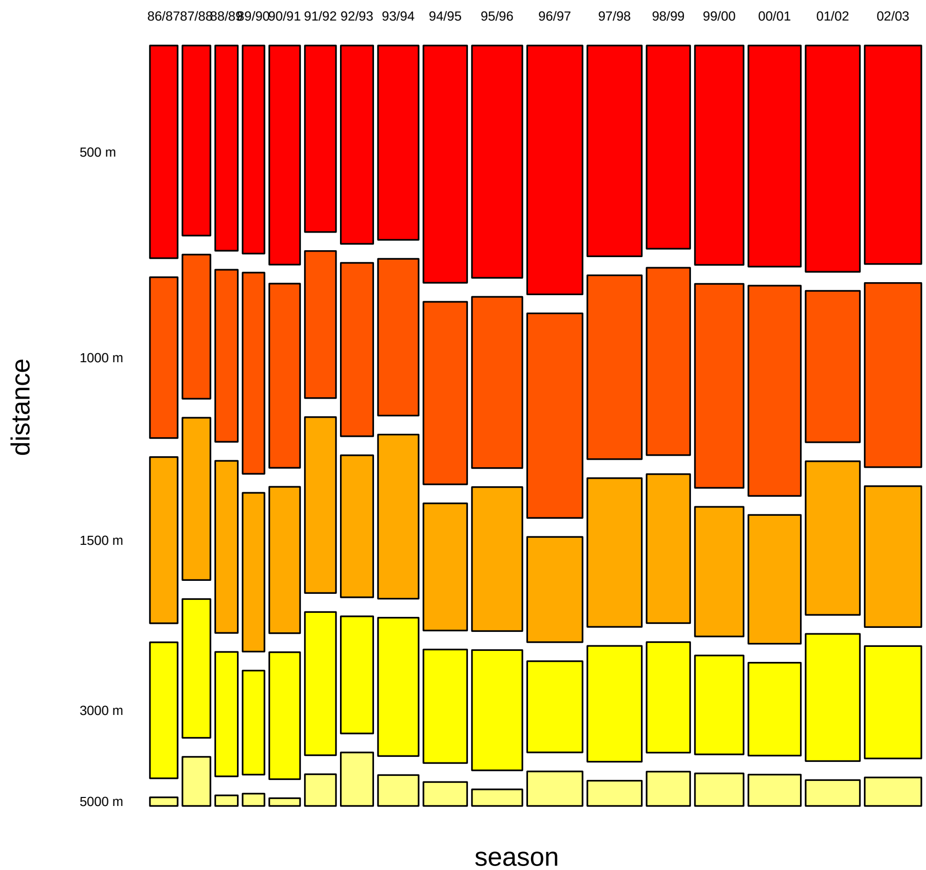


Figure 5: Distribution of observations by distance and season (women).

distance

n	missing	unique
21963	0	5

	500 m	1000 m	1500 m	5000 m	10000 m
Frequency	7371	5542	4534	3825	691
%	34	25	21	17	3

venue

n	missing	unique
21963	0	41

```
lowest : Albertville (FRA)     Asama (JPN)           Assen (NED)           Baselga di Pine (ITA) Bergen (NOR)
highest: Savalen (NOR)         Seoul (KOR)           Skien (NOR)           Warszawa (POL)        West Berlin (
```

home

n	missing	unique
21963	0	2

away result (19785, 90%), home result (2178, 10%)

season

n	missing	unique
21963	0	17

	86/87	87/88	88/89	89/90	90/91	91/92	92/93	93/94	94/95	95/96	96/97	97/98
Frequency	1102	1213	1020	1180	1011	1202	1241	1174	1331	1456	1479	1475
%	5	6	5	5	5	5	6	5	6	7	7	7

	98/99	99/00	00/01	01/02	02/03
Frequency	1248	1363	1396	1560	1512
%	6	6	6	7	7

alt.covered

n	missing	unique
21963	0	6

high altitude covered (2599, 12%), high altitude uncovered (2164, 10%)
lowland covered (7700, 35%), lowland uncovered (2656, 12%)
medium altitude covered (817, 4%), medium altitude uncovered (6027, 27%)

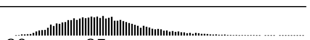
type2

n	missing	unique
21963	0	3

WC (20145, 92%), OG (964, 4%), WSDCh (854, 4%)

Women
7 Variables 15905 Observations

time.500m



n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
15905	0	4882	42.6	38.96	39.52	40.71	42.31	44.08	46.06	47.39

```
lowest : 36.91 36.98 37.03 37.06 37.07
highest: 54.59 54.63 54.73 54.96 54.98
```

distance

n	missing	unique
15905	0	5

	500 m	1000 m	1500 m	3000 m	5000 m
Frequency	5031	4173	3358	2699	644
%	32	26	21	17	4

venue

n	missing	unique
15905	0	38

```
lowest : Albertville (FRA)     Asama (JPN)           Assen (NED)           Baselga di Pine (ITA) Berlin (GDR)
highest: Salt Lake City (USA)  Seoul (KOR)           Skien (NOR)           Warszawa (POL)        West Berlin (
```

home

n	missing	unique
15905	0	2

away result (14079, 89%), home result (1826, 11%)

---

season

n	missing	unique
15905	0	17

	86/87	87/88	88/89	89/90	90/91	91/92	92/93	93/94	94/95	95/96	96/97	97/98
Frequency	634	641	517	500	706	712	742	930	1001	1155	1268	1247
%	4	4	3	3	4	4	5	6	6	7	8	8

	98/99	99/00	00/01	01/02	02/03
Frequency	997	1114	1201	1243	1297
%	6	7	8	8	8

---

alt.covered

n	missing	unique
15905	0	6

high altitude covered (2214, 14%), high altitude uncovered (1373, 9%)
lowland covered (6343, 40%), lowland uncovered (1620, 10%)
medium altitude covered (694, 4%), medium altitude uncovered (3661, 23%)

---

type2

n	missing	unique
15905	0	3

WC (14275, 90%), OG (797, 5%), WSDCh (833, 5%)

B Rink-specific fixed effects

Table 2: Fixed rink effects estimation results for men and women.

	Men			Women		
	estimate	st.err.	exp(est.)	estimate	st.err.	exp(est.)
Albertville (FRA)	3.6967	0.0021	40.313	3.7803	0.0029	43.830
Asama (JPN)	3.6842	0.0025	39.813	3.7668	0.0033	43.243
Assen (NED)	3.7030	0.0024	40.570	3.8030	0.0033	44.838
Baselga di Pine (ITA)	3.6912	0.0018	40.094	3.7835	0.0025	43.970
Bergen (NOR)	3.7184	0.0034	41.199			
Berlin (GDR)	3.7058	0.0020	40.683	3.7826	0.0025	43.930
Berlin (GER)	3.6862	0.0017	39.893	3.7771	0.0024	43.689
Butte (USA)	3.6817	0.0019	39.714	3.7754	0.0025	43.617
Calgary (CAN)	3.6587	0.0017	38.813	3.7466	0.0024	42.378
Chuncheon (KOR)	3.7105	0.0025	40.874	3.7956	0.0033	44.506
Collalbo (ITA)	3.6797	0.0020	39.636	3.7655	0.0028	43.184
Davos (SUI)	3.6868	0.0017	39.916	3.7895	0.0024	44.235
Den Haag (NED)	3.6947	0.0022	40.233	3.7820	0.0027	43.904
East Berlin (GDR)				3.7788	0.0033	43.762
Erfurt (GER)	3.6815	0.0022	39.707	3.7656	0.0028	43.188
Eskilstuna (SWE)	3.7177	0.0028	41.171			
Goteborg (SWE)	3.7463	0.0026	42.365			
Groningen (NED)				3.8194	0.0037	45.576
Hamar (NOR)	3.6712	0.0017	39.299	3.7590	0.0025	42.907
Harbin (CHN)	3.6809	0.0025	39.683	3.7726	0.0031	43.494
Heerenveen (NED)	3.6764	0.0017	39.503	3.7642	0.0023	43.128
Helsinki (FIN)	3.7179	0.0020	41.178	3.8078	0.0024	45.052
Ikaho (JPN)	3.6833	0.0025	39.776	3.7760	0.0031	43.639
Innsbruck (AUT)	3.6994	0.0017	40.424	3.7881	0.0025	44.173
Inzell (FRG)	3.6842	0.0019	39.813	3.7760	0.0025	43.641
Inzell (GER)	3.6893	0.0017	40.016	3.7830	0.0024	43.946
Jeonju (KOR)	3.7051	0.0024	40.655	3.8031	0.0030	44.842
Karuizawa (JPN)	3.6953	0.0020	40.257	3.7872	0.0027	44.134
Lake Placid (USA)	3.7085	0.0031	40.792	3.8068	0.0029	45.005
Larvik (NOR)				3.8008	0.0042	44.738
Medeo (KAZ)	3.7000	0.0027	40.449	3.7925	0.0037	44.367
Milwaukee (USA)	3.6690	0.0019	39.212	3.7627	0.0026	43.066
Nagano (JPN)	3.6772	0.0018	39.537	3.7678	0.0025	43.284
Obihiro (JPN)	3.7028	0.0022	40.561	3.7957	0.0028	44.510
Oslo (NOR)	3.7135	0.0020	40.997	3.8265	0.0029	45.903
Ostersund (SWE)	3.7073	0.0025	40.743			

Table 2: *(continued)*

	estimate	st.err.	exp(est.)	estimate	st.err.	exp(est.)
Roseville (USA)	3.7087	0.0020	40.801	3.7998	0.0026	44.692
Sainte-Foy (CAN)	3.7420	0.0031	42.183			
Salt Lake City (USA)	3.6555	0.0019	38.687	3.7429	0.0025	42.218
Savalen (NOR)	3.6963	0.0023	40.296			
Seoul (KOR)	3.7001	0.0019	40.451	3.7960	0.0026	44.522
Skien (NOR)	3.7121	0.0039	40.941	3.8074	0.0039	45.035
Warszawa (POL)	3.7117	0.0021	40.925	3.8050	0.0026	44.927
West Berlin (FRG)	3.7068	0.0019	40.722	3.8010	0.0030	44.744